

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2 0 0 4 年 3 月 1 7 日

出 願 番 号

Application Number:

特 願 2 0 0 4 - 0 7 6 4 7 2

パリ条約による外国への出願  
に用いる優先権の主張の基礎  
となる出願の国コードと出願  
番号

The country code and number  
of your priority application,  
to be used for filing abroad  
under the Paris Convention, is

J P 2 0 0 4 - 0 7 6 4 7 2

出 願 人

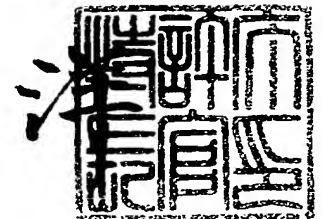
Applicant(s):

松下電器産業株式会社

2 0 0 5 年 4 月 2 7 日

特許庁長官  
Commissioner,  
Japan Patent Office

小 川



BEST AVAILABLE COPY

【官 報 号】	特 許 願
【整理番号】	2037460017
【提出日】	平成16年 3月17日
【あて先】	特許庁長官殿
【国際特許分類】	G06F
【発明者】	
【住所又は居所】	大阪府門真市大字門真1006番地 松下電器産業株式会社内
【氏名】	長光 左千男
【発明者】	
【住所又は居所】	京都市左京区鞍馬二ノ瀬町101
【氏名】	美濃 導彦
【発明者】	
【住所又は居所】	福岡県大牟田市正山町51-1
【氏名】	山肩 洋子
【特許出願人】	
【識別番号】	000005821
【住所又は居所】	大阪府門真市大字門真1006番地
【氏名又は名称】	松下電器産業株式会社
【代理人】	
【識別番号】	100067828
【弁理士】	
【氏名又は名称】	小谷 悦司
【選任した代理人】	
【識別番号】	100075409
【弁理士】	
【氏名又は名称】	植木 久一
【選任した代理人】	
【識別番号】	100109438
【弁理士】	
【氏名又は名称】	大月 伸介
【手数料の表示】	
【予納台帳番号】	012472
【納付金額】	21,000円
【提出物件の目録】	
【物件名】	特許請求の範囲 1
【物件名】	明細書 1
【物件名】	図面 1
【物件名】	要約書 1
【包括委任状番号】	0214505

【請求項 1】

人や物と操作の状況をセンシングするセンシング部と、  
人物特徴抽出部と操作特徴抽出部とからなる特徴抽出部と、  
操作前か後かを判断する状況判断部と、  
人や物を認識するための人物特徴テンプレートを有する人物認識部と、  
複数の人物認識結果において確信度を判断する人物確信度判断部と、  
操作を認識するための操作特徴テンプレートを有する操作認識部と、  
複数の操作認識結果において確信度を判断する操作確信度判断部と、  
人や物と操作の関連を認識するための人物操作特徴テンプレートを有する人物操作認識部と、  
複数の人物操作認識結果において確信度から判断する人物操作判断部と、  
判断結果を出力するための人物操作提示部とを備えた人物操作認識システム。

【請求項 2】

センシング部を、少なくとも画像センサと音センサにより形成した、  
請求項 1 に記載の人物操作認識システム。

【請求項 3】

人物操作提示部を、人物提示部、或いは、操作提示部の少なくともどちらかにより形成した、  
請求項 1、或いは請求項 2 に記載の人物操作認識システム。

【発明の名称】 人物操作認識システム

【技術分野】

【０００１】

本発明は、センサネットワークを基本にネット家電技術を応用した人の行動や物への操作を自動認識する人物操作認識システムに関するものである。

【背景技術】

【０００２】

従来の人物操作認識システムとしては、調理操作を事例に食材と操作を認識するものがあった（例えば、非特許文献１参照）。この場合には、光学カメラから得られる画像から食材の特徴量を知識テンプレート化して保有しておき、画像中における食材の確信度を求めている。ただし、得られる確信度はあまり高くは無く、食材のクローズドキャプションの情報などを付加して確信度の向上を図っている。

【０００３】

図３は、これら従来の人物操作認識システムの共通的な構成を示すものである。

【０００４】

図３において、観測デバイス１１０は、光学カメラ１１１やサーマルカメラ１１２、マイク１１３を用いて食材や調理操作の認識をしていた。特徴抽出部１２０における食材領域特徴抽出部１２１と操作特徴抽出部１２２とで調理の状況をセンシングし、状況判断部１３０は操作前か後かを判断し、食材特徴テンプレート１５１を参照し食材認識部１４１において対象である食材候補を列挙、食材確信度判断部１６１において確信度の高い食材を判断、食材提示部１９１で提示していた。一方調理操作の認識に関しては、操作特徴テンプレート１５２を参照し操作認識部１４２において対象である操作候補を列挙、操作確信度判断部１６２において確信度の高い順番に操作を判断、操作提示部１９２で提示していた。

【非特許文献１】 第１７回人工知能学会全国大会論文集、テキスト情報からの制約に基づく料理画像中の物体検出（no.2C1-02、June、2003）

【発明の開示】

【発明が解決しようとする課題】

【０００５】

しかしながら、前記従来の構成では、食材の認識と、調理の操作の認識を同時に関連付けずに、別々に判断しようとしていたので、食材と操作両方の確信度の認識が十分ではないという課題を有していた。

【０００６】

本発明は、前記従来の課題を解決するもので、人や物と操作の認識を高い確信度で実施できることを実現可能とした人物操作認識システムを提供することを目的とする。

【課題を解決するための手段】

【０００７】

前記従来の課題を解決するために、本発明の人物操作認識システムは、人や物と操作の関連を認識するための人物操作特徴テンプレートを有する人物操作認識部と、複数の人物操作認識結果において確信度から判断する人物操作判断部とを有し、人の行動や物への操作の自動認識を行う。

【０００８】

本構成によって、食材と操作の関連から確信度に基づく認識を実施することができる。

【発明の効果】

【０００９】

本発明の人物操作認識システムによれば、食材と操作両方に関して高い確信度の認識を実施することができる。

【発明を実施するための最良の形態】

【００１０】

以下、本発明の実施の形態について、図面を参照しながら説明する。

#### 【0011】

##### （実施の形態1）

図1は、本発明の実施の形態1における人物操作認識システムにおける構成図である。

#### 【0012】

図1において、センシング部10は各種カメラ等の画像センサやマイク等の音センサ等からなり、特徴抽出部20において人物特徴抽出部21や操作特徴抽出部22においてセンシング部10からの各種データから特徴量を抽出するものであり、状況判断部30において各操作の前後を判断し、もし操作の前の場合には人物特徴テンプレート51に基づき人や物を認識する人物認識部41と、人物確信度判断部61においてデータの確信度を定める。一方、操作の方も、操作特徴テンプレート52に基づき操作を認識する操作認識部42と、操作確信度判断部62においてデータの確信度を定める。また、操作後の人物特徴抽出部21からのデータも含めて、人物操作認識部71において、人物操作特徴テンプレート72を設けており、人と物や操作の有機的な関連から人物操作判断部80において、人や物と操作の認識結果を得て、人物操作提示部90とから構成している。人物操作提示部90には人物のみ、或いは操作のみを個別に提示する事も可能な構成としている（実施の形態1の構成）。

#### 【0013】

かかる構成によれば、人や物と操作の関連を認識するための人物操作特徴テンプレートを有する人物操作認識部と、複数の人物操作認識結果において確信度から判断する人物操作判断部とを有することにより、人の行動や物への操作の有機的な関連性に基づく自動認識を行うことが可能となり、高い確信度の人と物と操作の認識を実施することができる。

#### 【0014】

##### （実施の形態2）

図2は、本発明の実施の形態1である人物操作認識システムを、調理シーンへ具体的に適用した構成を示した図である。図2において、図3と同じ構成要素については同じ符号を用い、説明を省略する。

#### 【0015】

図2において、食材操作認識部271と食材操作特徴テンプレート272、食材操作判断部280から構成している。なお、食材操作提示部290は、食材提示部191と操作提示部192とから構成している。

#### 【0016】

かかる構成によれば、調理時における食材と調理操作の関連を認識するための食材操作特徴テンプレート272を有する食材操作認識部271と、複数の食材操作認識結果において確信度から判断する食材操作判断部280とを有することにより、調理の操作と食材の有機的な関連性に基づく自動認識を行うことが可能となり、高い確信度の食材と操作の認識を実施することができる。その結果、例えば図4に示したように、メニュー履歴に基づき食事管理をしてくれるシステムや、或いは図5のようにリアルタイムに音声対話などを使用して動的な調理支援タスクを実行するロボット家電のようなシステムも実現可能である。

#### 【0017】

ここから、具体的なシステム構築を行った事例を紹介する。特に、ユーザがテキスト教材に従い調理や裁縫などの生産加工作業を行う状況を想定し、カメラやセンサによる観測データに基づき、その生産加工活動を認識・支援するシステムの実現を目指した。そのためには、ユーザの扱っている物体や操作を、テキスト教材中で使われている言語として認識する必要があるが、通常これは困難である。そこで本発明を用いて、物体に対する単一の操作からなる作業を認識の最小単位とすることにより、（1）物体名と操作名の組み合わせに対する制約と、（2）この作業の結果得られる生成物が持つべき特徴に対する制約を利用して、物体と操作とを同時に認識する手法を提案する。食材に対する単一の調理操作をタスクとして取り上げ、食材と調理操作の認識において、上で述べた二つの制約が有

知ることができ、レシピと方針により確認した。シベリムで大衣し、長竹と調理器具で半紙し、認識した場合に比べ、本発明を適用した方が、より正しい認識結果を得られることを実験により確認した。

#### 【0018】

以下、詳細について実施例に基づき、説明する。

#### 【0019】

人間は日常生活において、物体に対し変形・結合・分割などの操作を数回にわたって行い、目的物を生成するといった生産加工活動を数多く行っている。このうち調理や裁縫、折り紙、木工作業などは、『レシピ』や『作り方』などのテキスト教材を学習・会得し行うユーザが多い。このようなユーザに対しては、「今どの物体に対し何をしようとしているのか？（物体・操作）」「どのテキスト教材に従い作業しているのか？（最終目的）」をシステムが推定することができれば、そのテキスト教材を基にユーザの作業を支援することが可能となる。そこで本実施例では、ユーザがシステムの持つテキスト教材のうちどれかに従い生産加工活動を行っている状況を想定し、カメラやマイク、センサなどによりユーザの動作を観測して、ユーザがどの物体にどの操作を行ったかを認識することにより、できる限り早い段階でユーザが従っているテキスト教材を推定することを目標とした。

#### 【0020】

テキスト教材は通常言語で記述されているため、物体や操作をテキスト教材中で用いられている用語として認識する必要がある。言語とその意味する観測データ中の特徴量とを一对一で対応付けることは難しく、認識候補が唯一に特定できない場合があるため、物体や操作単独での認識は困難である。そこで、物体に対する単一の操作からなる作業を最小単位とし、物体と操作とを総合的に認識することを考える（以降、このような作業を「単位作業」と呼ぶ）。ここで、単位作業における可能な物体名と操作名の組み合わせは、どれかのテキスト教材のどこかに現れるものしか存在せず、また、ある単位作業の結果生成される物体は常に決まった特徴を持つ。よってこれらを、単位作業の認識結果が満たさなければならない制約とすることにより、物体と操作を同時に認識し、物体と操作を単独で認識する場合に比べ、より確かな認識結果を得る手法を提案した。

#### 【0021】

得られた各単位作業の認識結果を、ユーザの作業手順に従い順次繋ぎ合わせてゆくと、現時点までのユーザの作業の流れを知ることができる。ユーザはシステムの持つどれかのテキスト教材に従い動作しているため、ユーザの作業の流れと各テキスト教材を比較すれば、ユーザの従っているテキスト教材が推定でき、かつ単位作業認識の誤りを補正することが可能となる。

#### 【0022】

以上のように本システムは「単位作業の認識」、「作業の流れによるテキスト推定」、「単位作業認識結果の修正」の三つの処理からなるが、本稿ではこのうち単位作業の認識の実現を目指し、それ以外はその枠組みについてのみ述べることにする。本実施例で対象としている生産加工作業の理解に関する従来例としては、音声やマニュアルなどのテキスト文解析によるアプローチや、部品の構造や設計順序を解析するアプローチがある。前者のアプローチでは、マニュアルやレシピを言語レベルで構造化した後、映像に付与されたクローズドキャプションや音声認識結果に対し、言語対言語での対応付けを行っている場合が多く、映像自体の認識はあまり行われていない。後者のアプローチでは、アセンブリ・プランニングやソフトウェア工学におけるプロジェクトプランニングの分野で研究が行われて来た。しかしこれらの目的は、効率のよい組立て手順や部品デザイン、モジュール分割の生成であり、言語など表現のモダリティが違ふシンボルとの関係については触れていなかった。よって、仮にマニュアルなど手順を示した情報があつたとしても、これを知識として利用することは考慮されていない。

#### 【0023】

以降、物体・操作・目的を同時に認識するための生産加工作業認識手法を提案、生産加工作業の具体的タスクとして調理を取り上げレシピを推定するためには物体や操作をどの

性反の祖てし認識する必女があるが、レシとガ例の和木を至に識調、調理におけるピ」で使った単位作業を取り上げ、1420種類のレシビと、実際の食材の観測データを用いて、提案した手法が有効であることを示した。最後に、実環境下で食材に対し包丁を使った調理操作を加える様子をカメラとマイクにより観測し、食材や調理操作を単独で認識した結果に比べ、本手法を使った方が良い結果が得られることを確認したので記述した。

#### 【0024】

まず、生産加工作業を、物体に対する単一の操作からなる単位作業と、それを繋ぎ合わせた作業全体の二段階に分け、まず単位作業を認識する手法について述べる。その後、作業全体の流れによるテキスト教材の推定と、それによる各単位作業認識結果の修正について、その枠組みを述べる。物体や操作をテキスト教材中の用語で認識するためには、予めすべての物体名や操作名を、それを観測した際に得られるべき特徴量にそれぞれ関連付けするのが一般的である。裁縫を例にあげると、「生地」や「糸」、「ボタン」などの物体名は『色情報』や『形状』などの特徴量に、「切る」「縫う」などの操作名は、ハサミの動きやミシンの操作情報などの特徴量にそれぞれ関連付けする。このとき、言語と特徴量では表現のモダリティが異なるため、言語空間で異なる用語を常に違う特徴量に割り当てることができるような特徴量空間を見付けるのは難しく、テキスト教材を推定するに十分な物体・操作の認識性能を得ることは困難である。しかしそもそも単位作業とは、物体と操作から、新たな特性を持つ物体を生成するまでが一つの作業であると考え、物体や操作といった『原因』のみならず、その『結果』である生成物も、単位作業認識に不可欠な要素であると言える。これは物体と操作を単独で認識したのでは利用できない性質である。そこで物体と操作を単位作業ごとにまとめて認識することを考える。単位作業を、実世界におけるインスタンス、観測データにおける特徴量、テキスト教材における言語の3つのレベルで表現したときの、互いの関係を図6に示す。単位作業には物体名と操作名からなるラベルが付けられており、単位作業が認識できれば物体名と操作名が決まる。単位作業単位で認識することにより利用可能となる制約は図6中の制約1、制約2であり、それぞれ以下のようなものである。

#### 【0025】

(制約1) 物体名と操作名の組み合わせは、どれかのテキスト教材のどこかに書かれているもののしか存在しない。

#### 【0026】

(制約2) 同じ単位作業よる生成物は、常に決まった何らかの特徴を持っているなければならない。

#### 【0027】

制約1の例としては、どのテキスト教材を見ても「ボタン」に対し「切る」操作を加えるといった記述はないため、「ボタン」と「切る」の認識結果が得られた場合、そのどちらかが間違っていると判定することができる。すなわち、どのテキスト教材中にも現れない物体名と操作名の組み合わせを導く認識候補は削除できる。

#### 【0028】

次に制約2の例としては、「ボタン」と「布」に「縫う」の操作を行った場合、「ボタン」は「布」の一部となり、必ず一緒に動くようになるため、もし独立に動くようであればその認識結果は間違っていると判定することができる。これは、「縫う」という操作が「物体同士を結合する」という性質をもっているため、いくら「ボタン」と「布」を近付け、その付近で「縫う」作業を行ったところで、両者が結合しなければ、「ボタン」と「布」を「縫う」ことにはならないためである。このとき、生成物に何らかの物体名を割り当て、生成物も単独で認識し、物体名・操作名・生成物名の可能な組み合わせを制約1とすればよいと思われるかもしれないが、生成物がある単位作業の結果であるために持つべき特徴は、生成物のもつ多くの性質の一部であるため、このような手法は適していない。このように、物体と操作を一つの単位作業として見ると、生成物はその単位作業の結果としてふさわしい特徴を持っているかということも、単位作業を認識する上で重要な制約となる。以上の2つの制約を各単位作業ごとに図6に示すようなテンプレートとして用意

ることにより、可能な物体と操作の組み合わせによる認識モデル等を山分けすることが可能になる。

#### 【0029】

単位作業の認識ができれば、それらを原材料から順に繋ぎ合わせてゆくことにより、現時点までのユーザの手順のフローを知ることができる。これはシステムの持つどれかのテキスト教材の手順フローの一部に完全に一致するはずである。しかし実際には、物体と操作の認識結果が誤っており、たまたまそれが実際とは違う単位作業に一致してしまった場合は、単位作業の認識結果に誤りが生じ、ユーザの手順フローがどのテキスト教材とも一致しないという状況が生じる。そこでまず、ユーザの現時点までの手順フローと、システムの持つすべてのテキスト教材の手順フローとをマッチングし、最も適合する手順フローをユーザが従っているテキスト教材と推定する。次に、そのテキスト教材の手順フローと、ユーザの手順フローで、一致しない部分に相当する単位作業の認識結果が誤りであったと判定し、そこで起こるべき単位作業の認識結果を導くよう修正を加える。以上のように、単位作業認識が実現できれば、物体・操作・目的を同時に認識する手法の実現が可能となる。

#### 【0030】

以上で述べた生産加工作業認識を現実世界で行うに際し、具体的タスクとして次に、調理を取り上げ、本発明を調理認識に適用することの有用性・有効性を調査した。

#### 【0031】

調理は多くの家庭において必要不可欠な日常活動の一つであり、一人の調理者が複数の食材や調理機器を同時に扱う複雑な行為であるため、機器操作やアドバイスなどの支援が求められている。目的物、すなわちメニューを推定することにより、毎日の一家の食事履歴が獲得できるため、健康管理やメニューの提案、食材管理など幅広い応用期待でき、要求の高いアプリケーションであると言える。生産加工作業としての調理は、物体が食材に、操作が調理操作に、目的物が最終的に完成する料理一品にそれぞれ対応する。またテキスト教材はレシピに対応する。

#### 【0032】

レシピを認識する上で、どの程度の種類の食材や操作が認識できる必要があるかを調べるため、ネスレ社のレシピサイトに掲載されている全レシピ1420種類を分析した。レシピ中で使われている全1032品目の食材のうち、1割以上のレシピに使われている食材182品目から調味料を除いた138品目の食材に着目する。このときこれらの食材名は、残りの894種類の食材名に比べ、おおまかな分類による種類名になっている。最も多くのレシピで使われていた食材は玉ねぎで、337種類のレシピに含まれており、仮にその他の食材が1品目特定できると最大106種類、2品目特定できると最大15種類までレシピ候補が減少する。全体では、食材が2品目特定できた場合、レシピの候補が20種類以上残る食材の組み合わせは3781組中55組、さらに30種類以上残る食材の組み合わせは21組あった。食材が2品目特定できた場合、残るレシピ候補数がそれぞれ1～20種類であった組み合わせの数を図7に示した。これによると、使用する食材が2種類わかると、全3781組中1724の組み合わせについては、レシピを一意に特定できることが分かる。よって本実施例では、ここで注目した138種類の食材について認識できればよいとする。しかしこの138種類の食材の中にも、外見が似通ったものが多く存在し、依然としてその認識は困難である。

#### 【0033】

前節で分析に用いたレシピで使われていた、調理操作を表す用語1020語のうち、使用回数が多かった順に上位30語の調理操作用語（全動詞の出現回数の58.8%をカバー）を図8に示す。これによると、似た特徴を持つと思われる用語が数多く存在し、これらを見分けることは比較的困難であることが予想される。以上の分析結果により、レシピを推定するために、認識しなければならない食材の数は最低138種類と多く、操作においても微小な違いを見分けなければならないため、食材や操作を単独で認識することは困難であると言える。



次に、調理における単位作業の認識に対し、本発明がどの程度有効であるかを検討するため、前述した2つの制約について調べる。調理では、各食材に任意の調理操作が同確率で発生するのではなく、食材によって偏りがあると予想される。そこで、全レシピから食材名と調理操作名の関係を抽出した。まず、前述の分析に用いたのと同じレシピ1420種類に対し、調理手順の説明文に対して係り受け解析を行い、食材名に係る調理操作用語を選出した。観測データによって判別可能な操作名として「剥く・切る」「焼く・炒める」「煮る・茹でる」「揚げる」「その他」の5種類を選出し、全ての調理操作用語をそのどれかに人手で分類した。例えば「人参と玉ねぎをみじん切りにします。」といった文からは、食材名「玉ねぎ」と操作名「切る」、食材名「人参」と操作名「切る」の関係がそれぞれ抽出される。全レシピにおいて出現頻度が高かった食材のうち、「切る・剥く」の操作を加えられ得る食材として、玉ねぎ（調味量を除いた900程度の食材のうち1位）、人参（8位）、トマト（11位）の3種類を選出し、各操作名の出現確率を調べた結果を図9（a）に示す。これによると、特に玉ねぎは煮るより炒める確率が高く、トマトは逆に炒めるより煮る確率が高い。またニンジンも煮るも炒めるもほぼ同確率で生じることが分かる。また、より細かいカテゴリ分をした場合も偏りが生じるかを調べるため、「切る」に関する操作名を「粗く切る」「細かく切る」に人手で分類した。ただし、用語「切る」は粗く切るか細かく切るかの判別ができないため、そのまま1つのカテゴリとした。「切る」「粗く切る」「細かく切る」の出現確率を調べた結果を図9（b）に示す。これによると、特に玉ねぎは他の食材と比べ細かく切られる確率が高く、顕著な偏りがあった。以上のように、食材名によって加えられる調理操作の発生確率は大きく異なり、食材と調理操作の関係制約として有用であることが確認できた。

## 【0035】

調理における調理操作のうち「切る」の操作を取り上げ、これを食材に対し加える際、元の食材が生成物となる課程での特徴量変化について調査した。被験者1名に以下の2種類の調理操作を行ってもらった。まず、ピーマンを細かく切る操作に関しては、開始後90秒程で千切りを終え、みじん切りへと移行。147秒後頃切り終えてまないたの外へ段階的に移し、150秒後頃まないたの上に何も無い状態で終了。次に、カボチャを粗く切る操作に関して、開始後130秒程でカボチャが半分に割れ、170秒後程度で4等分にされた。その後ヘタなどをとり、220秒頃カボチャをまないたの上に並べた状態で終了。作業の様子を上方に設置したカメラにより撮影し、食材領域の色相・彩度と領域サイズを1秒おきに算出した。これらの特徴量を選んだ理由は、これらが食材と調理操作の種類に依存すると考えたためである。時間の経過によるピーマンとカボチャの色相・彩度と領域サイズの変化を図10、図11にそれぞれ示す。Secondは食材が観測領域に現れてからの経過時間（秒）を表す。色相・彩度はそれぞれ12段階のヒストグラムを算出し、それをプロットしたときの等高線を、値が顕著に現れた部分（色相は0～5、彩度は0～10）のみ切り出して掲載している。まず色相・彩度に注目すると、始めはピーマンとカボチャは似たヒストグラムを示しているが、切ることにより食材の内部の色が外面に現れはじめると、特にカボチャの色が大きく変化し、両者の間に明らかな差が生じることがわかる。また、ピーマンは細かく切られるため画面に写される面が増え、サイズが大きく増加するが、カボチャは粗く切るため面がそれほど増えず、サイズが変わらないことがわかる。以上の結果により、「ピーマン」「カボチャ」に対する「切る」の操作からなる単位作業では、生成物の特徴は食材と操作の双方に強く依存し、食材と操作の認識において生成物の特徴量を考慮することが有用であることが確認できた。

## 【0036】

前述した包丁を使った単位作業について、実環境下でも正しく動作することを確認するため、実験を行った。前述で分析に用いたものと同種のレシピ48種類について、その中でよく使われていた食材12種類（ピーマン、キュウリ、カボチャ、ウリ、ネギ、キャベツ、玉ねぎ、ジャガイモ、レンコン、シイタケ、ゴボウ、サトイモ）を選出し、その48種類のレシピで「皮を剥く」「粗く切る」「細かく切る」の各操作が行われているかを調

た。このとき得られたデータは、食材と調理操作の組み合わせについて、市販のデジタル画像素材集より、各食材の操作前後の画像を選び、食材領域を手動で抽出して、食材の調理操作前後の特徴量を求めた。特徴量は、食材領域の色相（20段階）と彩度（10段階）のヒストグラムである。操作に関する特徴量としては、包丁で切られる回数を用い、予備実験により各操作ごとに回数を設定した。以上の処理により、単位作業のテンプレートを作成した。被験者の前上方向にカメラを設置し、まないたを含む手周辺を光学およびサーマルカメラにより撮影した（図12）。通常、包丁を使った操作では食材の視覚的特徴が変化するため、光学カメラのみを使った食材の領域抽出およびトラッキングは困難である。しかし、食材は加工直前まで冷蔵庫や冷暗所（通常10度～15度）で保管されており常温より低温であるか、過熱されて高温である場合が多いため、サーマルカメラを食材領域抽出に用いることを考える。この処理の手順は以下になる。

#### 【0037】

1. サーマルカメラによる温度画像（図13、311）と、光学カメラによる光学画像（図13、321）のそれぞれについて背景差分画像（図13、312、322）を作成する。このとき背景画像は、各画素について過去一定時間（10分）で最も数多く出現した画素値を採用することにより作成する（図13、313、323）。

#### 【0038】

2. 各背景差分領域に対し、膨張処理、その後収縮処理を加えることにより、小さな穴や島のノイズを除去する（図13、314、324）。

#### 【0039】

3. 温度画像と光学画像の座標系を合わせるため座標変換し（図13、315、325）、積領域を獲得する（図13、316）。この領域をまとめごとに分割し（図13、317）、各食材領域を獲得する（図13、318）。

#### 【0040】

以上の処理手順により実際に食材領域抽出を行った結果、背景差分のしきい値などを調節することにより、ほとんどの場面で食材領域のみが抽出されることを確認した。このようにして得られた食材領域について、色相・彩度のヒストグラムを算出する。また、包丁で切られる回数を評価するため、被験者の前方にマイクを設置し、環境音を収録して、包丁がまないたとぶつかる衝突音を人手で書き出した。以上の処理により得られる各特徴量を観測データとする。システムは、観測データと各単位作業のテンプレートとの距離を算出し、これから各単位作業の認識結果としての確信度を算出する。

#### 【0041】

以上のシステムを実装し、「カボチャを粗く切る」作業の認識を行った。単位作業認識結果より算出した、食材および調理操作ごとの認識結果の確信度を各々図14（a）と図14（b）に示した。食材の外見だけの特徴を使い食材単独で認識した場合はウリの確信度が最も高かったが、本手法を適用した結果、カボチャの確信度が最も高くなり正しい認識結果が得られた。また、操作の認識結果は、操作単独で特徴量を比較した場合でも「粗く切る」と認識されていたものの、本手法を適用すると確信度がより上昇した。以上の結果により、単位作業として食材と操作を統合的に認識することにより、食材と操作を単独で認識した場合に比べ互いの認識結果が向上することを確認した。

#### 【0042】

本実施例では、ユーザがテキスト教材に従い行う生産加工活動の認識を目標とし、まず、単一の操作からなる単位作業を取り上げ、（1）物体名と操作名との可能な組み合わせに関する制約と、（2）生成物が満たすべき特徴量の制約を用いて、物体と操作を同時に認識する手法を提案した。それらの単位作業認識結果を得ることにより、作業全体の手順の流れを獲得し、ユーザの従っているテキストの推定と、各部分作業の認識結果の修正を行う枠組について述べた。具体的なタスクとして調理を取り上げ、レシピを推定する目的においては、食材や調理操作単独での認識は難しいことを確認した。調理での単位作業について、食材と操作、および生成物の間に、上で挙げた2つの制約が使えることをレシピ分析と映像データ解析により確認した。最後にカメラとマイクにより実際の調理を観測す

る大域でいい、大環境にしても物体と操作を半独立認識する物口に見え、必ず広て適用したほうがより正しい認識結果を導き出すことを確認した。

#### 【0043】

なお、操作後を対象とした人物テンプレートを準備して人物認識部や人物確信度判定部を構成することも考えられ、確信度がより向上すると予想できるが、現実的にはそのような構成は準備できないことは特筆すべきである。

#### 【0044】

また、本実施の形態において、調理を例に実施例を記述したが、図1のような構成において、人の行動を対象に、ネット家電等の機器を操作した際の人や機器の操作前後の状態変化に基づき人間行動を予測することにも使用することも可能であり、センシング部10にウェアラブルセンサを構成しても良いし、人物特徴テンプレート51として人の顔認証データや室温等の環境状態による温度変化する空気等も対象と考えられる。さらに、操作特徴テンプレート52として人の歩き方や表情を使用すれば、セキュリティや対話インタラクションのツールとしても使用できる。本実施の形態2は一例である。

#### 【0045】

ここで、用語の説明として、確信度について記述する。

#### 【0046】

人間やコンピュータが何らかの知識を持っているとき、それが事実かどうか分からない場合がある。このような知識のことを「信念(belief)」と呼ぶ。本発明では、認識器がある標本をあるカテゴリに属すと判断する際、その判断が事実であると信じる信念の度合を「確信度(belief factor)」と呼び、具体的には各カテゴリのプロトタイプと標本との距離などから算出する。例えば赤くて丸い物体が観測された場合、ほぼリンゴだと思うがトマトである可能性も捨てがたいと判定したい場合、観測物体とリンゴ・トマトのプロトタイプデータとの類似度をもとに、リンゴの確信度を0.8、トマトの確信度を0.2というように表現することができる。用語はこれに限定する必要は無い。

#### 【産業上の利用可能性】

#### 【0047】

本発明にかかる人物操作認識システムは、人物操作特徴テンプレートに基づく人物操作認識部を有し、人や物と操作を自動認識するシステム一般として有用である。また、ユーザがテキスト教材に従い調理や裁縫などの生産加工作業を行う状況を想定し、カメラやセンサによる観測データに基づき、その生産加工活動を認識・支援するシステムとして、調理以外にも裁縫、折り紙、木工作業等の用途にも応用できる。

#### 【図面の簡単な説明】

#### 【0048】

【図1】本発明の実施の形態1における人物操作認識システムの構成図である。

【図2】本発明の実施の形態2における人物操作認識システムの構成図である。

【図3】従来の人物操作認識システムの構成図である。

【図4】本発明の実施の形態2における人物操作認識システムの応用事例1の図である。

【図5】本発明の実施の形態2における人物操作認識システムの応用事例2の図である。

【図6】本発明の実施の形態2における単位作業における制約の図である。

【図7】本発明の実施の形態2におけるレシピ候補数の図である。

【図8】本発明の実施の形態2における上位30語の調理操作名を示す図である。

【図9】本発明の実施の形態2における食材名における操作名の発生確率の図である。

【図10】本発明の実施の形態2における「ピーマンを細かく切る」操作の特徴量変化の図である。

【図11】本発明の実施の形態2における「カボチャを細かく切る」操作の特徴量変化の図である。

【図 1 2】本発明の実施の形態 2 における具体的なシステム構成図である。

【図 1 3】本発明の実施の形態 2 における具体的なシステムフロー図である。

【図 1 4】本発明の実施の形態 2 における食材と調理操作認識の実験結果を示す図である。

【符号の説明】

【0 0 4 9】

7 1 人物操作認識部

7 2 人物操作テンプレート

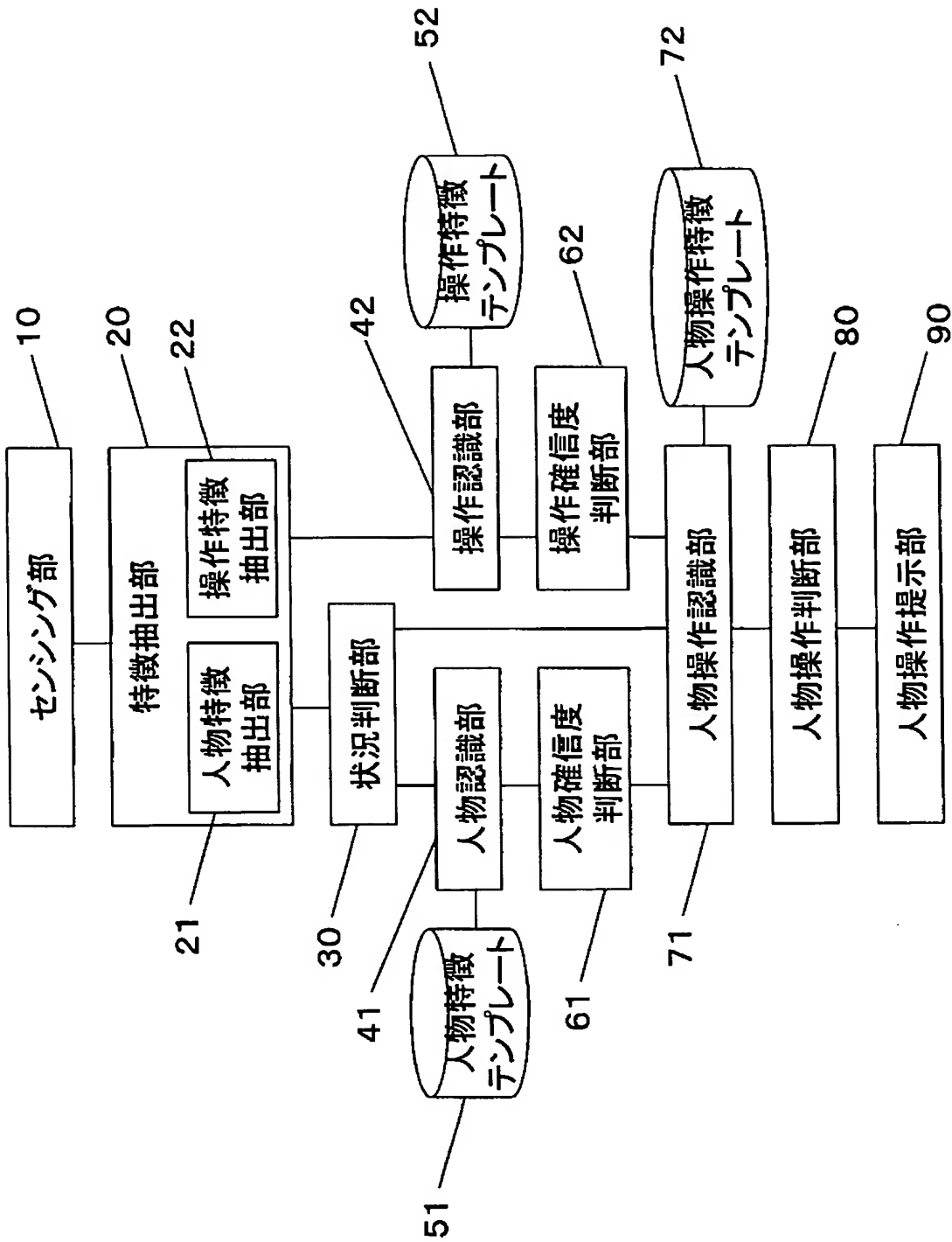
8 0 人物操作判断部

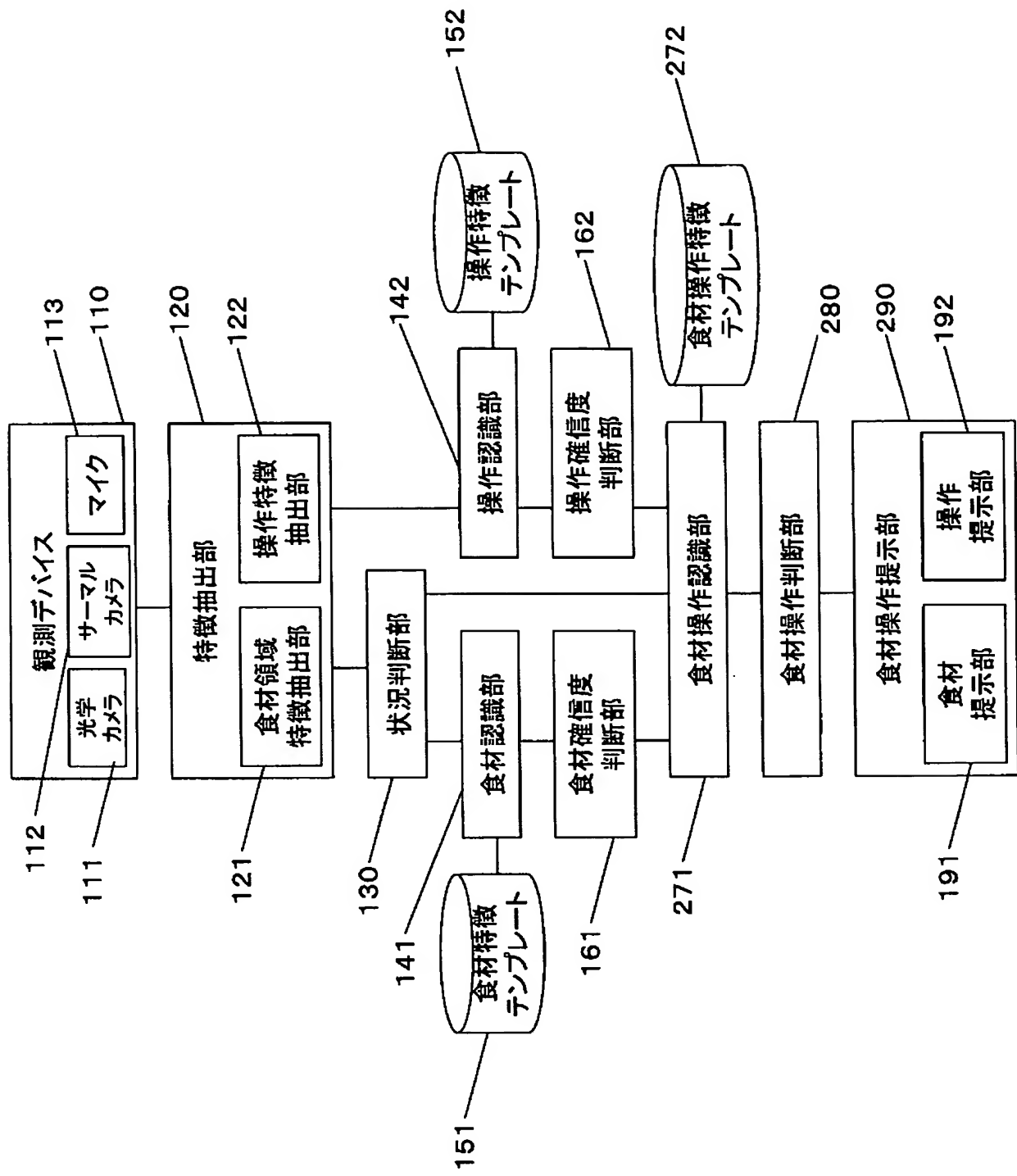
1 1 1 光学カメラ

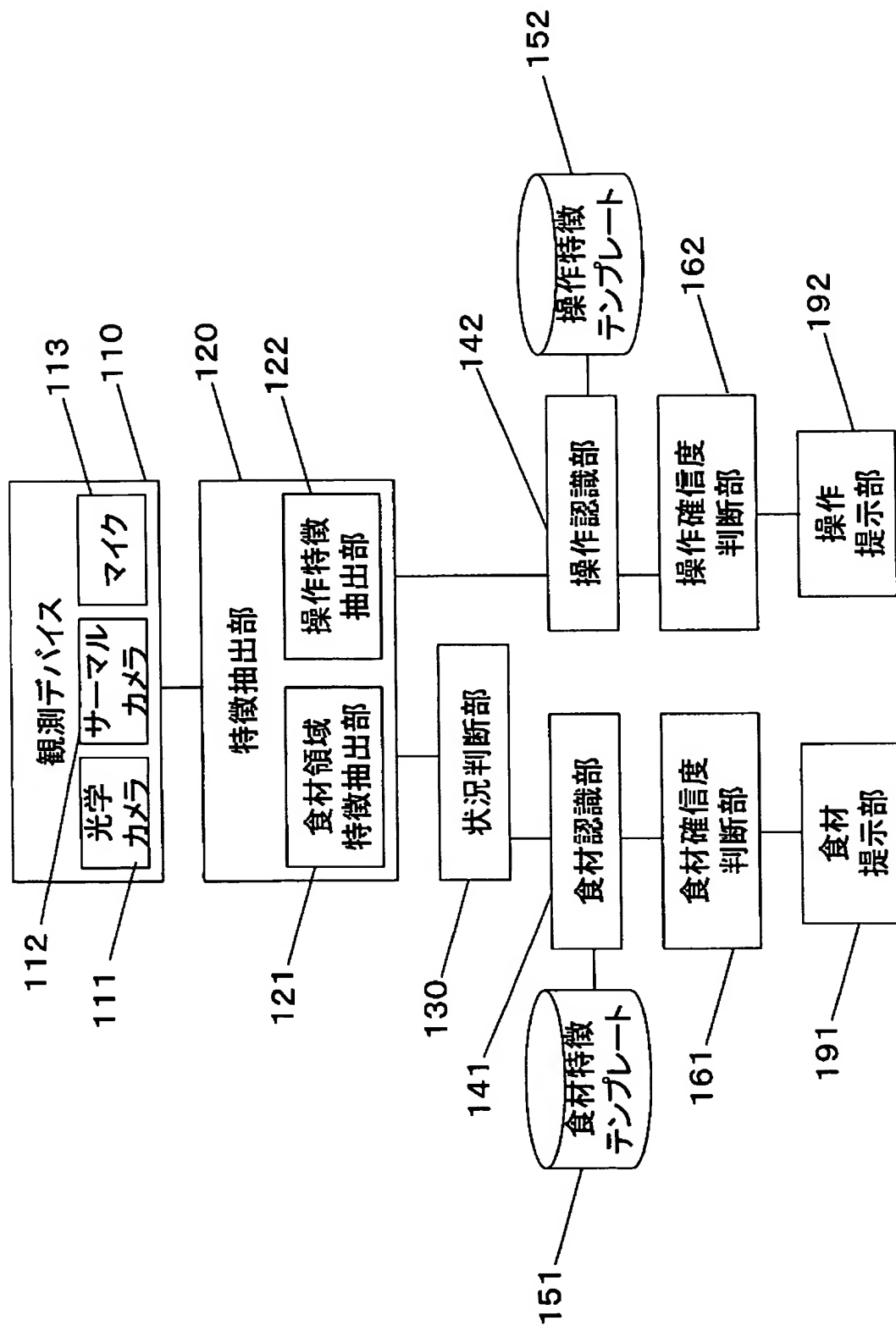
1 1 2 サーマルカメラ

1 1 3 マイク

【図 1】







## アプリケーション1:

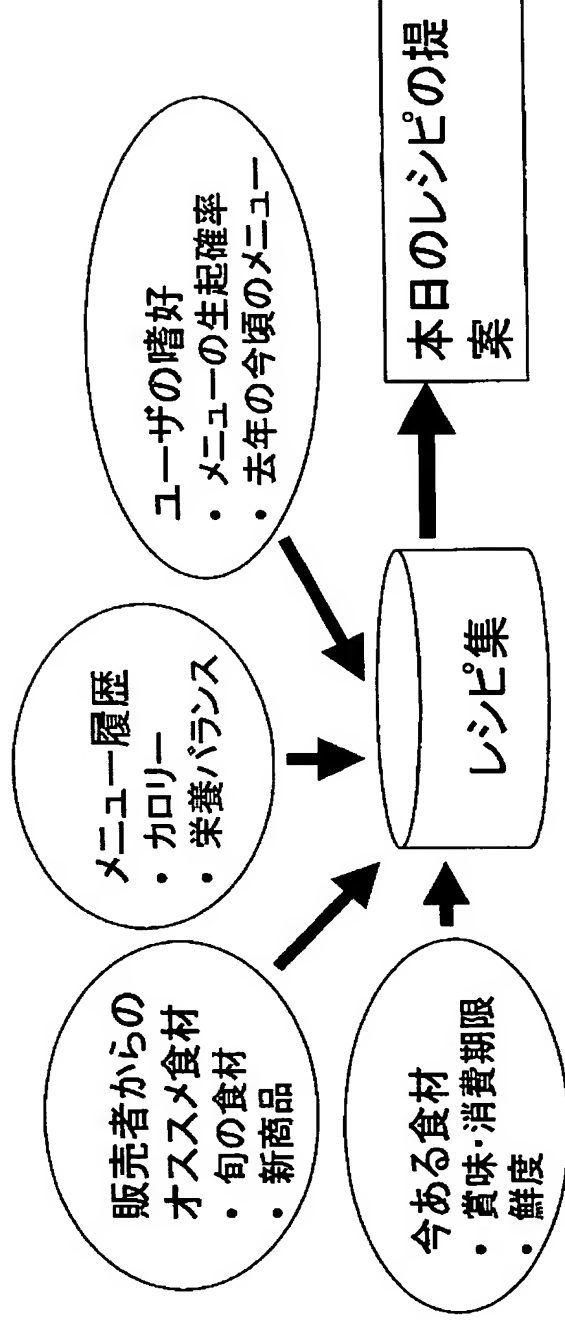
# メニュー履歴にもとづく食事管理システム

医療費増額が社会問題に

⇒ 健康な人を不健康にしない食事管理の必要性

しかし、自身の健康を疑わない人に食事管理を課すのは困難

🏠 ・ユーザに意識させることなく調理履歴を獲得、適切なアドバイスをを行う





## アプリケーション2:

# 音声対話による動的な調理支援タスク



ユーザの調理進行にあわせ音声対話によりアドバイス

- ・ 現在わかっていることでアドバイスを生成
- ・ わからないことは聞く⇒現在わかっていることからユーザの

発話を予測

音声認識候補

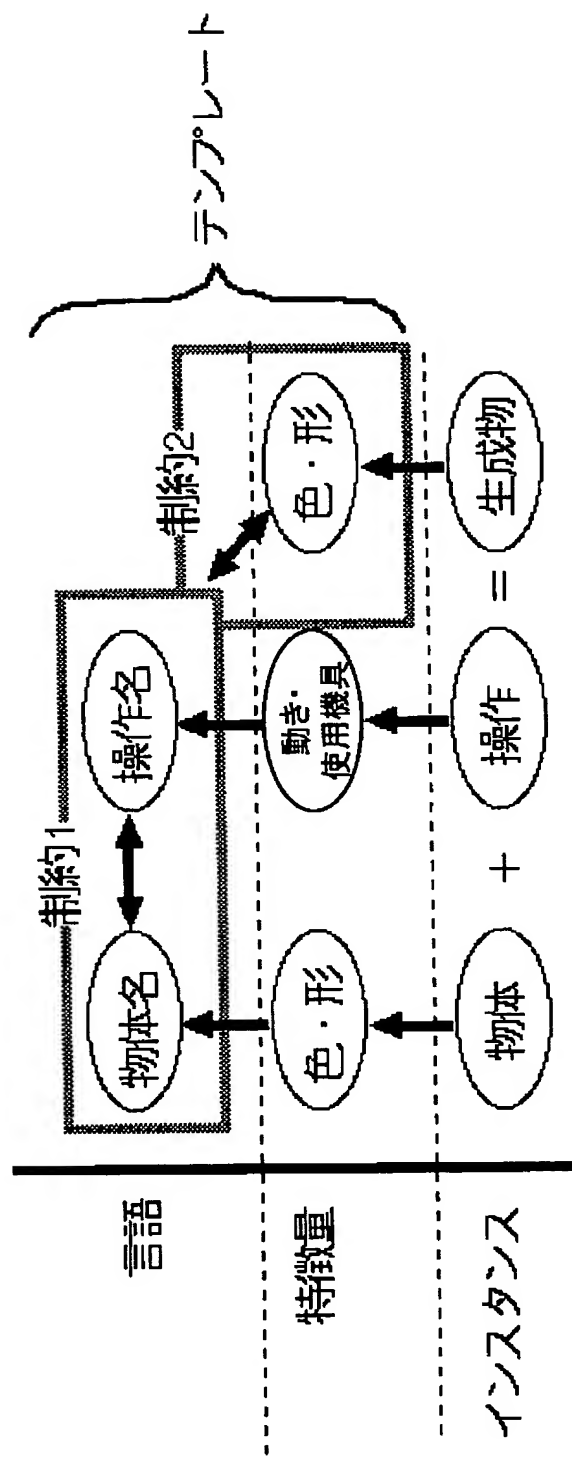
「シーフードスパゲティが、  
カニの刺身が、  
蛸をゆめるとは」

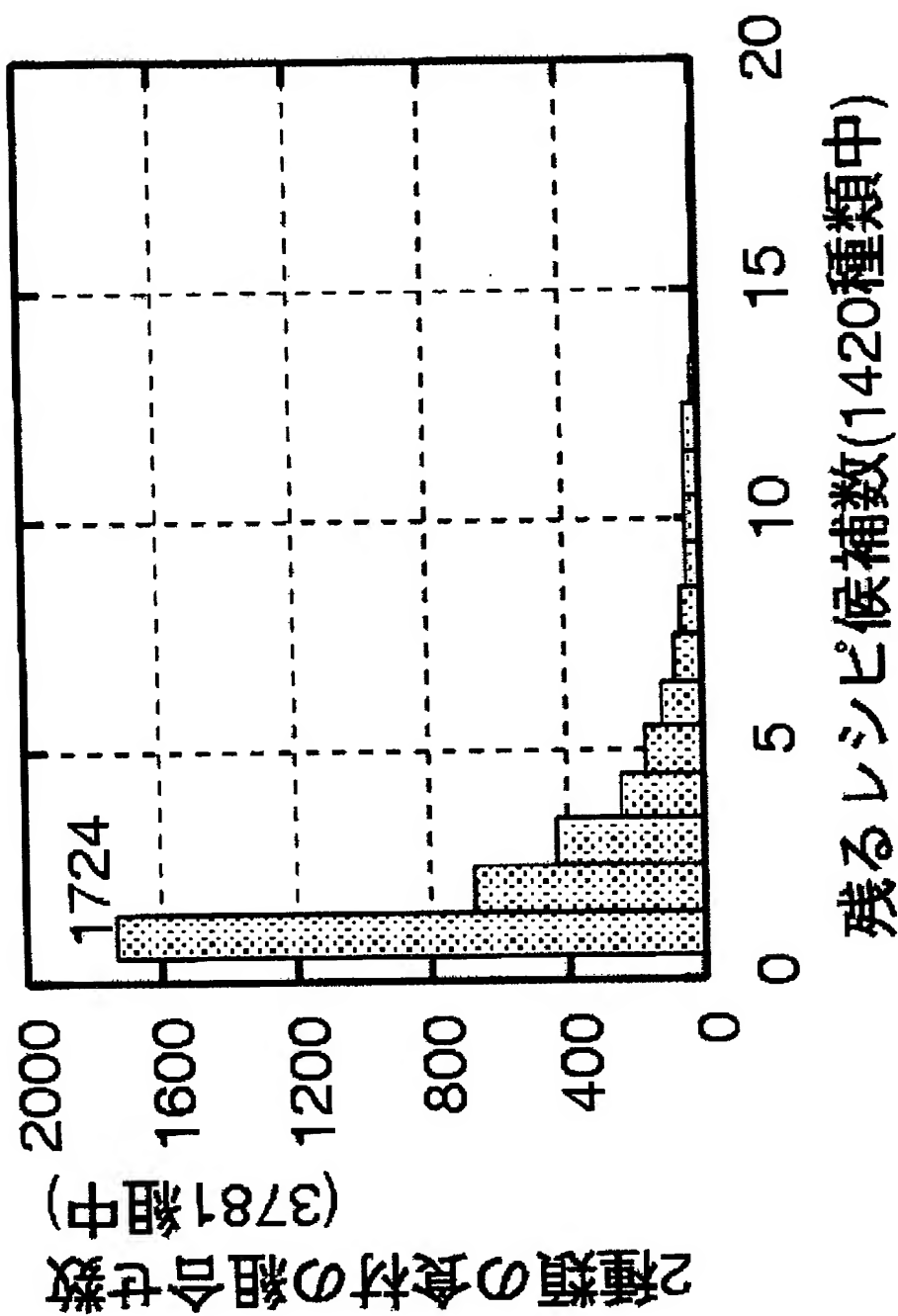
何を作るんですか？

シーフード  
スパゲティ

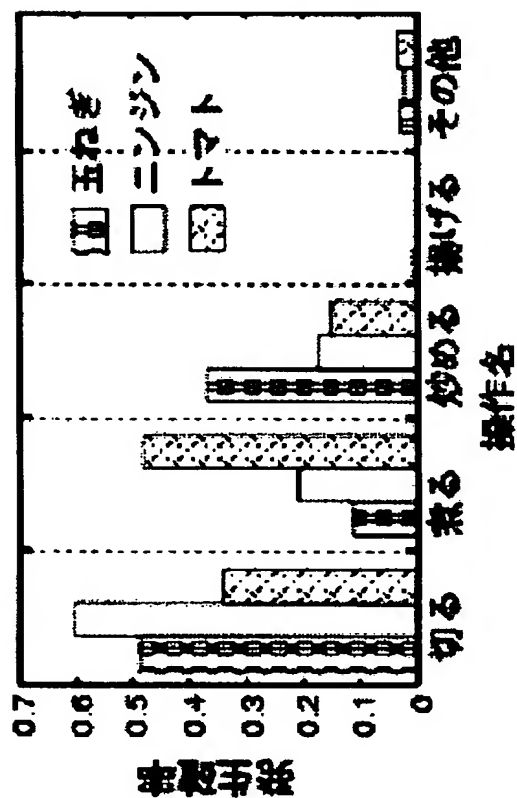
まず胴の部分に指を入れて  
裏返して内臓を取り出します。

蛸を軟らかくするにはもう  
少し煮たほうがいいですよ。

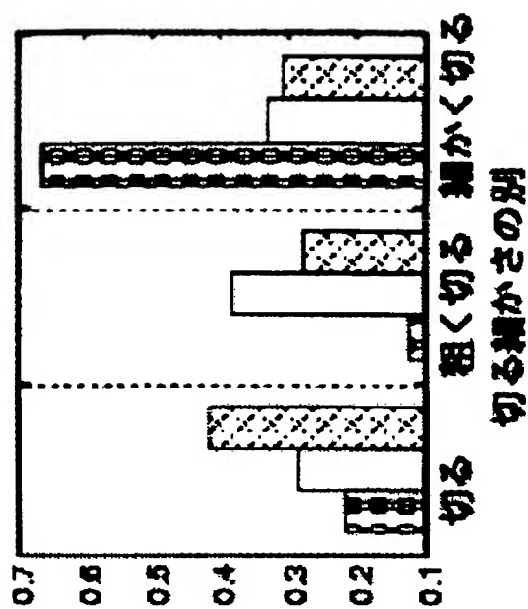




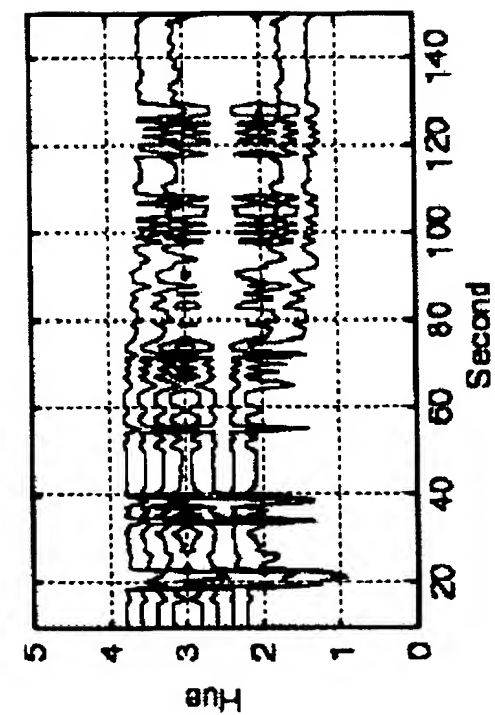
分類	出現頻度 (30語中)	用語
混ぜる	46. 2%	加える, 混ぜ合わせる, 注ぐ, 合わせる, 和える, まぶす, 流し入れる, 泡立てる
焼く・炒める	12. 3%	炒める, 焼く, 炒め合わせる
煮る・茹でる	11. 8%	煮る, 煮立てる, 煮込む, 茹でる, ひと煮立ちさせる
切る・剥く	8. 25%	切る, みじん切りにする
揚げる	1. 1%	揚げる
その他	20. 4%	調味する, 盛る, 熱す, 添える, 飾る, 溶かす, 調味, 冷ます, 冷やす, 蓋をする



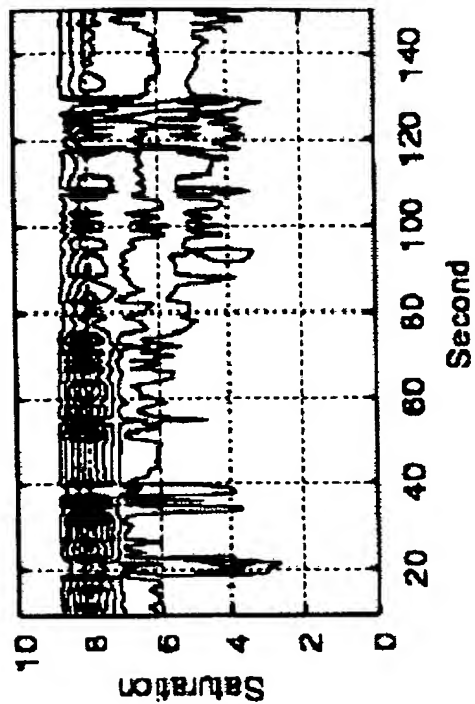
(a)



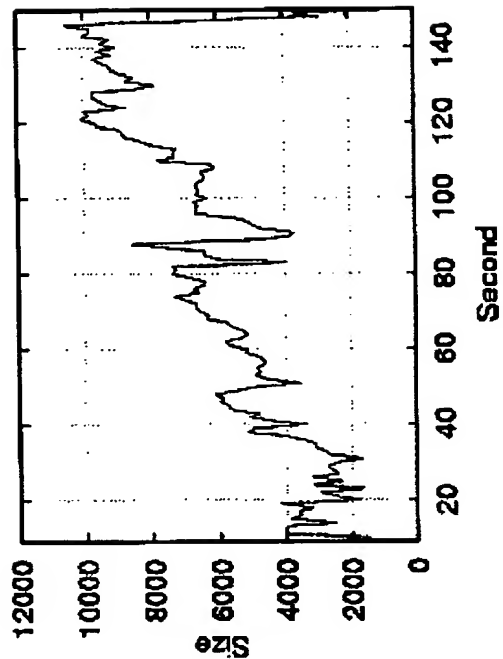
(b)



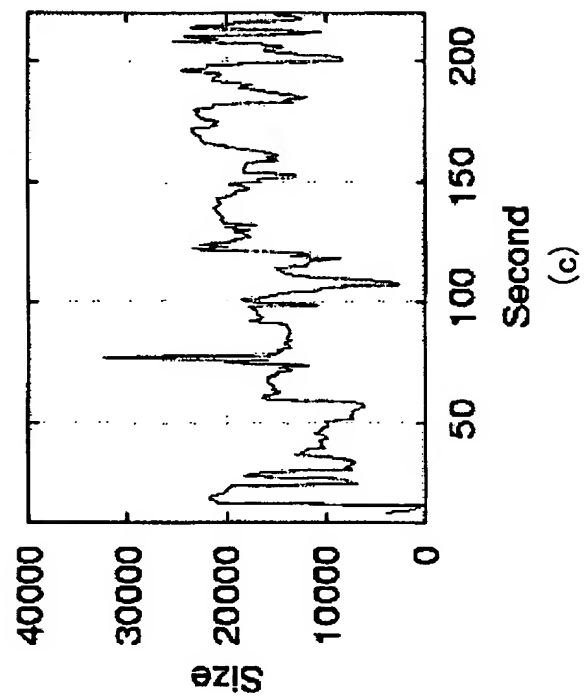
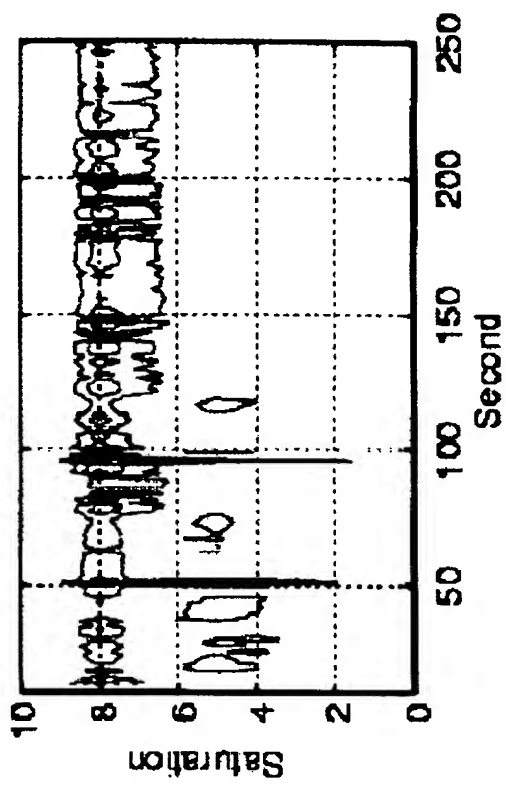
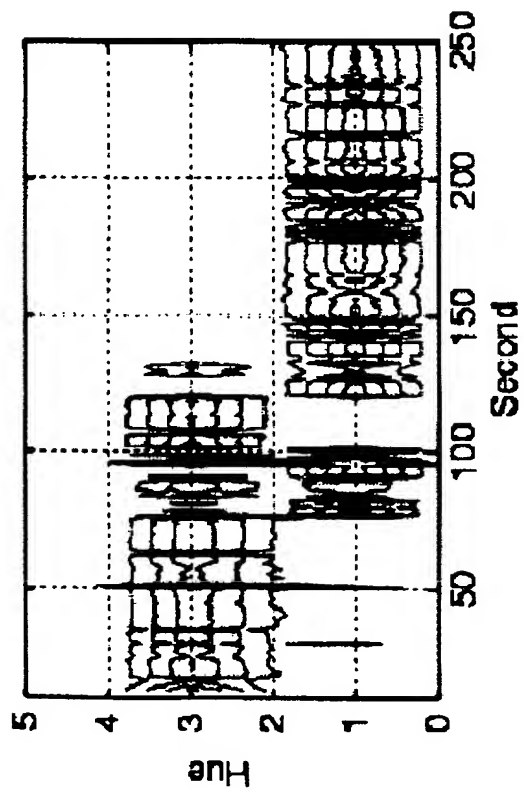
(a)

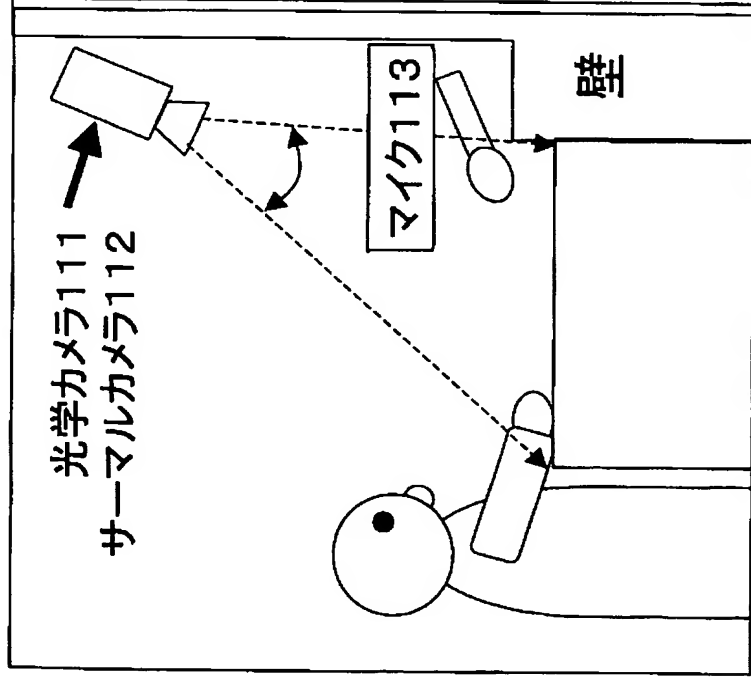


(b)

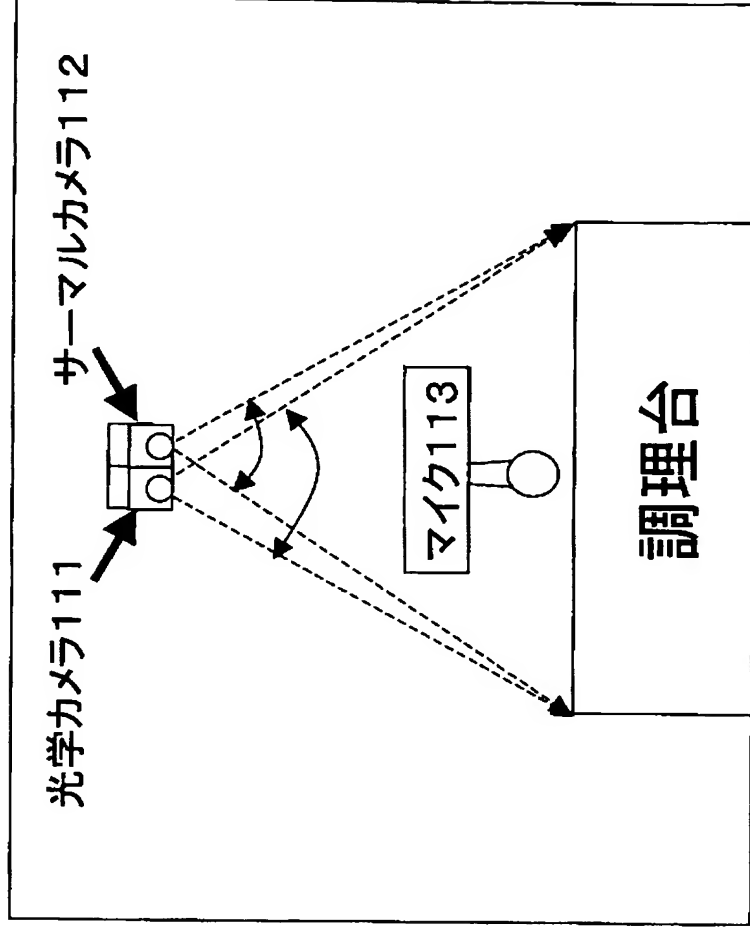


(c)



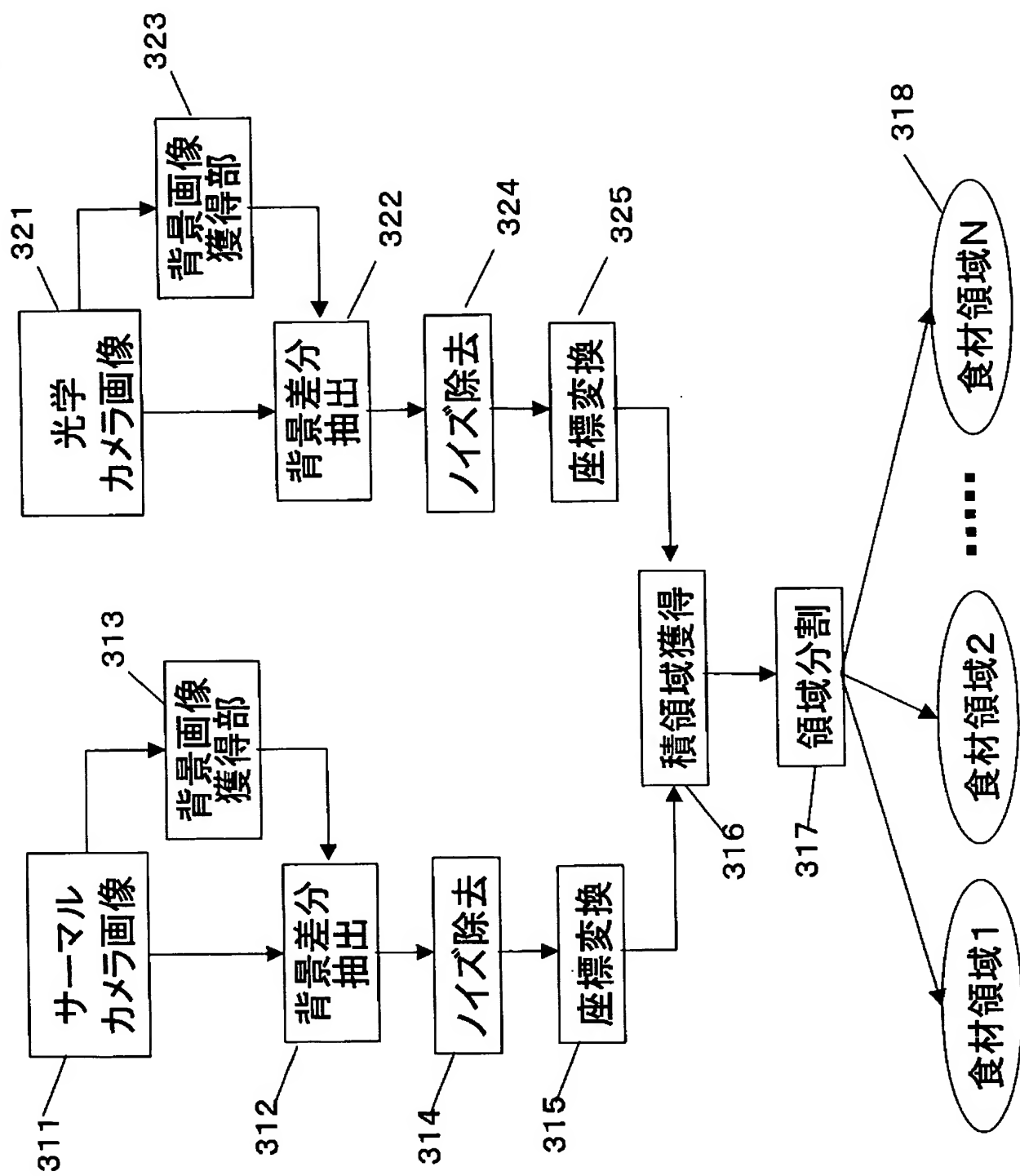


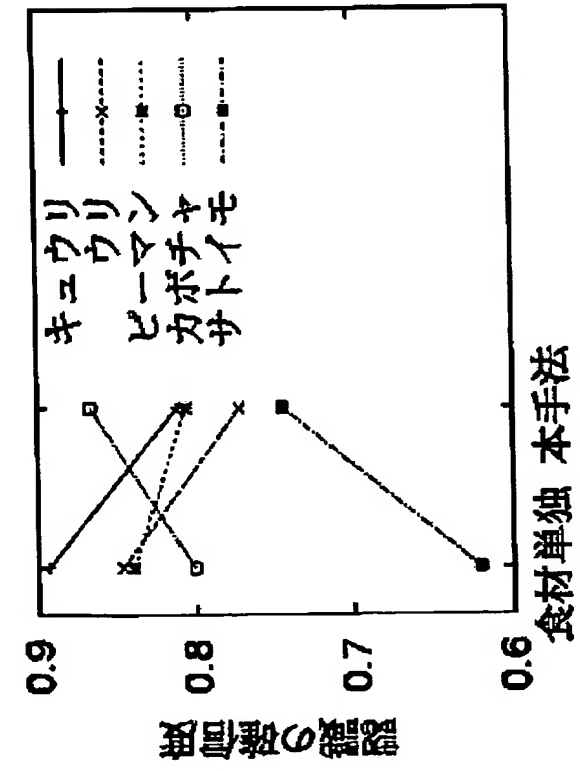
(a)



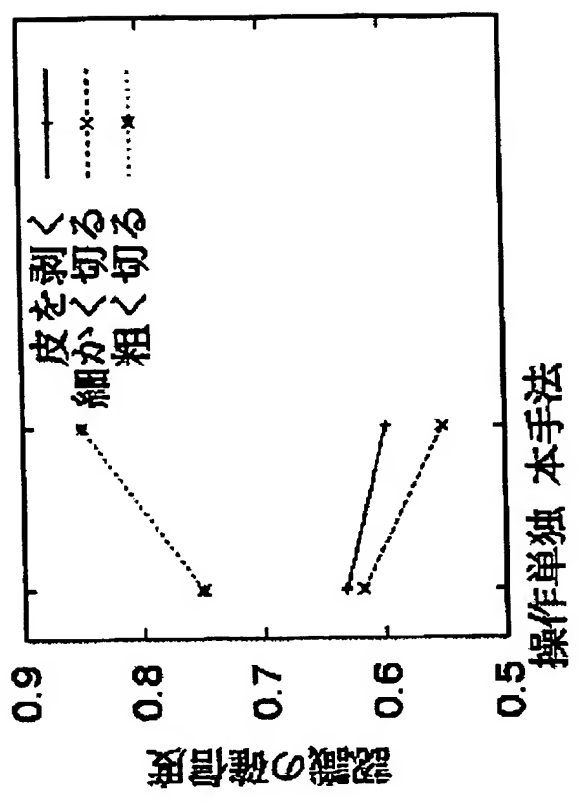
(b)







(a)



(b)

【要約】

【課題】 人や物と操作の認識を高い確信度で実施できることを実現可能とすること。

【解決手段】 人や物と操作の関連を認識するための人物操作特徴テンプレート72と、人物操作認識部71と、複数の人物操作認識結果において確信度から判断する人物操作判断部80とを備え、食材と操作の有機的な関連から確信度に基づく認識を実施することにより、食材と操作両方に関して高い確信度の認識を実施することができ、人の行動や物への操作の自動認識が可能となる。

【選択図】 図1

,

0 0 0 0 0 5 8 2 1

’ 19900828

新規登録

大阪府門真市大字門真 1 0 0 6 番地

松下電器産業株式会社

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP05/004828

International filing date: 17 March 2005 (17.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: JP  
Number: 2004-076472  
Filing date: 17 March 2004 (17.03.2004)

Date of receipt at the International Bureau: 20 May 2005 (20.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**